

# Aportaciones metodológicas para luchar contra el fraude: un reto multidisciplinar

## Methodological contributions for the fightind fraud: A multidisciplinary challenge

Antonia Ferrer-Sapena y Enrique A. Sánchez-Pérez

**Ferrer-Sapena, Antonia; Sánchez-Pérez, Enrique A.** (2018). "Aportaciones metodológicas para luchar contra el fraude: un reto multidisciplinar". *Anuario ThinkEPI*, v. 12, pp. 356-360.

<https://doi.org/10.3145/thinkepi.2018.60>

Publicado en *IweTel* el 30 de noviembre de 2017



**Resumen:** Estado de la cuestión sobre métodos científicos utilizados contra el fraude financiero, principalmente de tipo técnico (informáticos, económicos, matemáticos). Aunque algunas de las áreas científicas involucradas no están relacionadas con la tecnología -por ejemplo, la sociología-, es importante insistir en que todas ellas proporcionan herramientas útiles para detectar el fraude. La detección y prevención del fraude financiero es una tarea multidisciplinar, por lo que la solución a este problema de urgente actualidad deberán aportarla equipos multidisciplinarios.

**Palabras clave:** Fraude financiero; Técnicas científicas; Detección; Multidisciplinar; Triángulo del fraude.

**Abstract:** State of the art scientific methods can be used against financial fraud, mainly of a technical nature (computer, economic, mathematical). Although some of the scientific areas involved are not related to technology, such as sociology, it is important that all disciplines provide useful tools to detect fraud. The detection and prevention of financial fraud is a multidisciplinary task, so the solution to this urgent problem needs to be provided by multidisciplinary teams.

**Keywords:** Financial fraud; Scientific techniques; Detection; Multidisciplinary; Fraud triangle.

## 1. Introducción

En los últimos años, la irrupción del periodismo de datos en el contexto del análisis de la evasión fiscal y otras formas de fraude ha abierto nuevas vías para entender la lucha contra los delitos económicos. La aparición en los medios de comunicación de los llamados *Papeles del paraíso* es un nuevo episodio de una forma de hacer públicas ciertas actividades económicas ilegales –o al menos, escasamente éticas- de personas, entidades y empresas, continuación de una serie de escándalos entre los que cabe destacar la *Lista Falsiani* o los *Papeles de Panamá*.

La característica principal que hace que los procesos que hay detrás de estas noticias sean especiales, es que se basan en la gestión analítica de una cantidad ingente de datos. Datos económicos, balances, información contractual, relaciones sospechosas... Todos estos datos, en formatos diversos, aparecen en documentos que muchas veces contienen información numérica, aunque generalmente es textual, y de la que pueden extraerse conclusiones sobre participaciones en negocios y relaciones cuestionables entre personas y entidades.

Así pues, la gestión de información masiva y su interpretación es la base de estos procesos de análisis, y el problema principal al que se enfrentan los analistas consiste en determinar cuál debe ser la metodología adecuada para abordar estas investigaciones. El objeto de este tipo de análisis es el que tradicionalmente se consideraba una auditoría de cuentas, o un estudio legal de las obligaciones fiscales. Sin embargo, la cantidad de datos hace simplemente imposible que un grupo de expertos considere todos los documentos, y extraiga de ellos la información relevante. Debe ser, por tanto, la tecnología, la que permita hacer una primera criba para facilitar una información estructurada, sobre la cual se pueda hacer la necesaria investigación. Atendiendo a los aspectos técnicos del problema, la pregunta que surge es: ¿cómo puede hacerse este tratamiento inicial de los datos?, y ¿qué disciplinas pueden proporcionar herramientas para llevarlo a buen puerto?

**“La característica principal que hace que los procesos que hay detrás de estas noticias sean especiales, es que se basan en la gestión analítica de una cantidad ingente de datos”**

Se presenta aquí una panorámica de cuáles son los elementos conceptuales y técnicos que se han utilizado para el análisis de los casos de fraude. Se trata sólo de un examen superficial, en el que con seguridad omitiremos contribuciones importantes, y en el que además no abordaremos las contribuciones necesarias desde el punto de vista de la legislación y la gestión tributaria, por ejemplo. Nos limitaremos a temas más relacionados con los aspectos técnicos y conceptuales más cercanos a disciplinas de carácter académico. Debe tenerse en cuenta que las técnicas que se desarrollarán sin duda en los próximos años para abordar estos problemas, no se restringirán a la investigación de casos como los comentados, de una dimensión enorme, sino que permitirán la investigación de casos de fraude a menor escala –como el desvío de dinero en la contratación de

un cierto servicio por parte de una empresa pública-, hasta convertirse en herramientas de uso cotidiano, como pasa por ejemplo con las aplicaciones de tipo *business intelligence* en el campo empresarial. Sin embargo, estas situaciones comparten con los grandes casos comentados antes la característica central de la necesidad de gestionar una gran cantidad de información.

## 2. Contribuciones metodológicas

En los últimos años, y en paralelo a la aparición de nuevas técnicas analíticas y nuevos casos de fraude a gran escala, algunos grupos de trabajo han comenzado a estudiar cuáles son los aspectos metodológicos de las técnicas que hasta el momento se han aplicado (**Richhariya; Singh, 2012; Trompeter et al., 2014**). Podemos clasificar estas contribuciones según su origen en tres grandes grupos:

a) Cuando se piensa en el fraude en términos abstractos, los contextos metodológicos sugeridos en primera instancia son los que facilitan los estudios sociológicos, económicos y legales. Obviamente, los aspectos económicos son fundamentales, y las leyes deben adaptarse al nuevo entorno tecnológico, pero la aplicación necesaria de estas disciplinas debe hacerse después del análisis inicial, dirigido a aportar una estructura. La sociología sí que nos aporta un contexto conceptual inicial atractivo desde el que es posible entender a grandes rasgos los hechos constitutivos de fraude. La teoría del triángulo del fraude, establecida hace más de 50 años (ver **Beasley, 1980; Deter, 1988; Martin, 2015; Dorminey et al., 2012**), establece que la comisión de cualquier fraude (sea del tipo que



Figura 1

sea: económico, científico o académico) responde a la concurrencia de tres situaciones:

- la presión que las circunstancias ejercen sobre la persona que tomará la decisión de cometerlo;
- la oportunidad de convertir esa posibilidad en un acto;
- la racionalización, que permite la justificación ante sí mismo de la realización de un acto calificable, al menos, de deshonesto.

La figura 1 muestra las relaciones entre los distintos elementos del modelo.

A efectos prácticos, no parece que esta teoría facilite técnicas concretas, pero sí el contexto metodológico y conceptual necesario. Se pueden encontrar en la literatura científica estudios de prevención del fraude en las administraciones públicas, de carácter estadístico y basados por ejemplo en encuestas de opinión realizadas en grupos de inspectores fiscales y miembros de la administración, cuyos diseños se basan en la teoría del triángulo del fraude (Naruedomkul; Rodwanna; Wonglimpiyarat, 2010). Otras contribuciones de carácter puramente técnico tienen también su origen en modelos de base sociológica asociados a la cuantificación del riesgo de fraude (Mock; Srivastava; Wright, 2016).

**“Parece necesario el establecimiento de una teoría central que unifique los diferentes argumentos y técnicas matemáticas usadas, y que permita entender todos los procedimientos y argumentos en un contexto común”**

b) El segundo gran bloque de posibles contribuciones metodológicas son las que podríamos denominar de carácter informático. En este grupo debemos incluir el tratamiento de los datos y su interpretación desde el punto de vista estrictamente formal, que aporte una cierta homogeneidad y estructura a la información. El problema es cómo convertir un conjunto de documentos de texto, de tipo mercantil, correos electrónicos, balances económicos, contratos, etcétera, en un material formado por elementos matemáticos claros, principalmente índices numéricos y relaciones entre elementos. Para ello se pueden utilizar programas de interpretación semántica de texto, de inteligencia artificial, algoritmos de extracción de información contable y otras herramientas muy especializadas. El resultado debe ser una base de datos no estructurada, con unas propiedades directamente condicionadas por el contexto matemático

elegido para el tratamiento posterior, y que comentaremos en la siguiente sección.

c) Los métodos automáticos de búsqueda de información en internet –incluyendo las llamadas técnicas de *scrapeo* o de minería de datos, entre otras-, también juegan un papel importante, puesto que permiten la obtención de datos que pueden estar relacionados directamente con el problema tratado o facilitar una descripción del contexto, también necesaria para dar una respuesta adecuada (Ngai et al., 2011). Por ejemplo, para estudiar el historial de contratación de una cierta empresa por parte de otra compañía o de una administración pública, es imprescindible disponer previamente de información sobre sus contratos anteriores, cambios de nombre o posibles impagos. Existen muchas herramientas informáticas disponibles, y la adecuación de cada una de ellas depende del problema concreto.

### 3. Aspectos matemáticos

Los métodos matemáticos utilizados hasta la fecha han sido diversos, y normalmente adaptados a contextos concretos. Por ejemplo, para la detección del fraude en el uso de tarjetas de crédito se utilizan programas de detección numérica de “puntos lejanos” (*outliers*) dentro de un cierto conjunto (Zhao et al., 2016). El contexto natural para el análisis numérico es el de los métodos estadísticos y de las matemáticas relacionales. Sin embargo, estos problemas requieren de la consideración de cada aspecto específico desde un punto de vista apropiado, sobre todo para su comprensión en un contexto conceptual que permita su interpretación y posterior resolución. Por ejemplo, según se trate de la determinación de desequilibrios anómalos en la competencia entre empresas, o de la investigación sobre la existencia de relaciones sospechosas en la concesión de contratos, se puede recurrir a la teoría de juegos o a la teoría de grafos, respectivamente (Wilks; Zimbelman, 2004; Szárnyas et al., 2016). Es posible encontrar en la bibliografía científica algunos artículos que presentan una clasificación de las técnicas matemáticas usadas hasta el momento en la lucha contra el fraude, y en qué aspectos concretos de esta actividad se han usado –*neural networks*, *support vector machine*, *decision trees*, y otros (West; Bhattacharya, 2016 y las referencias que aparecen en este trabajo)–. Existen programas informáticos específicos de análisis financiero basados en algunas de estas teorías, como el *Neo4j*, basado en la estructuración de la información en forma de grafos enriquecidos, y que se ha convertido en un instrumento analítico habitual para el análisis de procesos financieros y del fraude. La organización de la información permite

representaciones gráficas como la que se puede observar en la figura, que facilita enormemente la investigación.

Sin embargo, parece necesario el establecimiento de una teoría central que unifique los argumentos y técnicas matemáticas usadas, y permita entender todos los procedimientos y argumentos en un contexto común. La evaluación del riesgo de fraude, la detección de delitos económicos y la previsión de situaciones que puedan favorecer los procesos fraudulentos, se realizarían así dentro de un mismo entorno conceptual, basado en la sociología -la teoría del triángulo del fraude y sus actualizaciones-, y de carácter formal. Esta teoría permitiría integrar diferentes métodos y argumentos, técnicas y procedimientos para el análisis del fraude en un marco común.

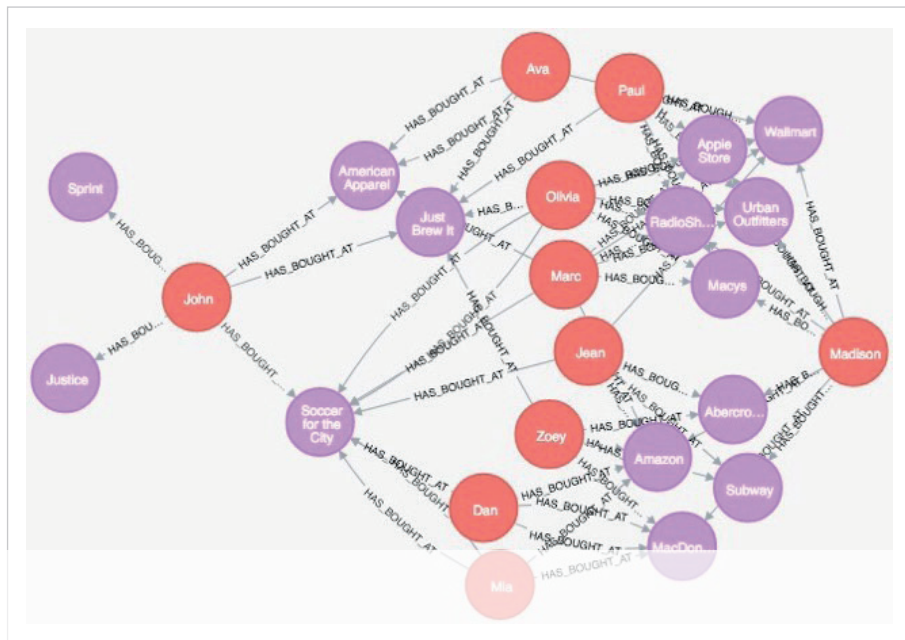


Figura 2

en el concurso convocado para su asignación, tiempos de pago, número de contratos por empresa,..., y las relaciones conocidas entre los elementos característicos de cada contrato –por ejemplo, relaciones probadas entre empresas-, añadiendo a cada relación establecida aquellos datos considerados relevantes, generalmente de carácter numérico. Toda esta información se puede estructurar en un grafo, en el que los nodos son los elementos de los contratos y las relaciones entre ellos. Este tipo de estructura de datos que permitan definir una métrica interna en el grafo. Esta métrica, que indica la proximidad relativa entre los elementos del grafo, es la base para la aplicación de las técnicas analíticas al problema, y puede permitir el agrupamiento de datos (*clustering*), la determinación de relaciones ocultas y de carácter probabilístico y otras acciones de organización y análisis de la información. La definición matemática de los entornos de los elementos –por ejemplo, vecindarios de un empresario importante y supuestamente corrupto-, y otros elementos topológicos, pueden proporcionar el marco formal conceptual requerido. Algunas ideas en esta dirección se pueden encontrar en Szárnyas et al. (2016).

Estas consideraciones indican la imprescindible concurrencia de diferentes disciplinas en la comprensión y la investigación de los delitos económicos: periodistas, sociólogos, juristas, economistas, informáticos, documentalistas, y matemáticos son sólo algunos de los profesionales que deberán entenderse y aportar su conocimiento para la necesaria comprensión global del fraude.

“Es imprescindible la concurrencia de diferentes disciplinas en la comprensión e investigación de los delitos económicos: periodistas, sociólogos, juristas, economistas, informáticos, documentalistas, y matemáticos, entre otros”

Aunque en la actualidad no existe todavía, al parecer, ningún contexto teórico formal que se pueda considerar hegemónico en el que entender el fraude, sí hay algunos indicios que parecen marcar hacia donde podrían dirigirse los pasos de los investigadores. Los datos en bruto no permiten una interpretación, de manera que es necesaria la construcción de estructuras relacionales cuantificadas, es decir, relaciones entre elementos -empresas, personas, facturas, administraciones, ...- a las que además se puedan añadir características y cuantificar su intensidad. Por ejemplo, el estudio de las contrataciones de una administración pública con empresas durante un año, empezaría por la estructuración de los elementos relevantes en cada contrato –cantidades pagadas, empresas participantes



#### 4. Conclusiones

La lucha contra el fraude, a todos los niveles, se presenta como un nuevo desafío social para los próximos años. Como normalmente sucede en los momentos en que aparece una nueva necesidad social, la respuesta tecnológica tiende a dar soluciones puntuales a problemas concretos. A pesar del importante esfuerzo que se está haciendo ya para aportar los conocimientos necesarios y adaptar los existentes a este contexto, todavía no existe un núcleo teórico que permita la creación de una nueva disciplina que dé solución a los problemas que el fraude plantea con una metodología clara. Todo parece indicar que la creación de estas estructuras conceptuales llevará un tiempo, y la respuesta final será netamente multidisciplinar. Desde un punto de vista formal, un contexto teórico que en la actualidad puede aportar esa deseada metodología única podría venir dado por la creación de bases de datos con estructura de grafo dirigido, junto con una topología adecuada, definida por una métrica. Esta distancia no tiene por qué ser necesariamente simétrica: la asimetría en su definición facilitaría la modelación de la falta de reciprocidad en las relaciones que se establecen en el grafo.

Los documentalistas y científicos de la información se presentan como actores imprescindibles en la resolución de este reto, entre otros profesionales, con los que será necesario formar grupos esencialmente multidisciplinarios. La formación, el interés y el trabajo de algunos grupos de técnicos en documentación deben pues dirigirse hacia este fin.

#### 5. Referencias

**Abdullahi, Rabi'u; Mansor, Noorhayati** (2015). "Fraud triangle theory and fraud diamond theory. Understanding the convergent and divergent for future research". *International journal of Academic Research in Accounting, Finance and Management Sciences*, v. 5, n. 4, pp. 38-45.

<http://www.iiste.org/Journals/index.php/EJBM/article/viewFile/26274/26919>

**Dorminey, Jack; Fleming, A. Scott; Kranacher, Mary-Jo; Riley, Richard A. Jr.** (2012). "The evolution of fraud theory". *Issues in accounting education*, v. 27, n. 2, pp. 555-579.

<https://goo.gl/BWgANR>

<https://doi.org/10.2308/iace-50131>

**Mock, Theodore; Srivastava, Rajendra; Wright, Arnold M.** (2017). "Fraud risk assessment using the fraud risk model as a decision aid". *Journal of emerging technologies in accounting*, v. 14, n. 1, pp. 37-56.

<https://goo.gl/ww28ss>

<https://doi.org/10.2308/jeta-51724>

**Naruedomkul, Pornchai; Rodwanna, Pannipa; Wonglimpiyarat, Jarunee** (2010). "Organization frauds in Thailand: A survey on risk factors". *International journal of criminal justice sciences*, v. 5, n. 1, pp. 203-219.

<https://goo.gl/Zb2rkq>

<https://doi.org/10.1016/j.dss.2010.08.006>

**Ngai, Eric W.T.; Hu, Yong; Wong, Y. H.; Chen, Yijun; Sun, Xin** (2011). "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature". *Decision support systems*, v. 50, n. 3, pp. 559-569.

<https://doi.org/10.1016/j.dss.2010.08.006>

**Richhariya, Pankaj; Singh, Prashant** (2012). "A survey on financial fraud detection methodologies". *International journal of computer applications*, v. 45, n. 22, pp. 975-1007.

<https://goo.gl/EZAuuu>

**Szárnay, Gábor; Koovári, Zsolt; Salánki, Ágnes; Varró, Dániel** (2016). "Towards the characterization of realistic models: Evaluation of multidisciplinary graph metrics". In: *Proceedings of the ACM/IEEE 19th International Conference on Model Driven Engineering Languages and Systems*, pp. 87-94.

<https://inf.mit.bme.helsinki.fi/publications/models2016-metrics.pdf>

**Trompeter, Gregory M.; Carpenter, Tina D.; Jones, Keith L.; Riley, Richard A. Jr.** (2014). "Insights for research and practice: What we learn about fraud from other disciplines". *Accounting horizons*, v. 28, n. 4, pp. 769-804.

<https://goo.gl/ArzYf5>

<https://doi.org/10.2308/acch-50816>

**West, Jarrod; Bhattacharya, Maumita** (2016). "Intelligent financial fraud detection: a comprehensive review". *Computers & security*, v. 57, pp. 47-66.

<https://goo.gl/KDzuof>

<https://doi.org/10.1016/j.cose.2015.09.005>

**Wilks, I. Jeffrey; Zimbelman, Mark F.** (2004). "Using game theory and strategic reasoning concepts to prevent and detect fraud". *Accounting horizons*, v. 18, n. 3, pp. 173-184.

<https://doi.org/10.2308/acch.2004.18.3.173>

**Zhao, Jie; Lau, Raymond Y.K.; Zhang, Wenping; Zhang, Kaihang; Chen, Xu; Tang, Deyu** (2016). "Extracting and reasoning about implicit behavioral evidences for detecting fraudulent online transactions in e-Commerce". *Decision support systems*, v. 86, pp. 109-121.

<https://goo.gl/P7BdmV>

<https://doi.org/10.1016/j.dss.2016.04.003>

**Antonia Ferrer-Sapena**

Universitat Politècnica de València

Escuela Técnica Superior de Informática (ETSINF)

<http://mugi.webs.upv.es>

<http://ctranspa.webs.upv.es>

[anfersa@upv.es](mailto:anfersa@upv.es)

**Enrique A. Sánchez-Pérez**

Universitat Politècnica de València

Instituto Universitario de Matemática Pura y

Aplicada (IUMPA)

[easancpe@mat.upv.es](mailto:easancpe@mat.upv.es)

Register for free at <https://www.scipedia.com> to download the version without the watermark